

WEST

 Generate Collection

L9: Entry 1 of 5

File: USPT

Oct 17, 2000

Filed: NOV 19, 97 (AK Date)

DOCUMENT-IDENTIFIER: US 6134532 A

TITLE: System and method for optimal adaptive matching of users to most relevant entity and information in real-time

DATE FILED (1):
19971114Abstract Text (1):

A system and method for selecting and presenting personally targeted entities such as advertising, coupons, products and information content, based on tracking observed behavior on a user-by-user basis and utilizing an adaptive vector space representation for both information and behavior. The system matches users to entities in a manner that improves with increased operation and observation of user behavior. User behavior and entities (ads, coupons, products) and information (text) are all represented as content vectors in a unified vector space. The system is based on an information representation called content vectors that utilizes a constrained self organization learning technique to learn the relationships between symbols (typically words in unstructured text). Users and entities are each represented as content vectors.

Brief Summary Text (3):

The invention relates to associating entities and information with behavior. More particularly, the invention relates to a system and a process for targeting and delivering advertising, coupons, products, or informational content to users based upon observed behavior.

Brief Summary Text (5):

The widespread availability of the World Wide Web (web) and Internet services has resulted in a unique set of advertising opportunities. Unlike conventional "broadcast" media, such as television and radio, the web is a "narrowcast" medium that allows the user to have higher levels of control over the information they receive. Since users can control the retrieval of information, the advertising techniques utilized in the conventional broadcast model has become less effective and can alienate potential customers as a result of the "shotgun" effect. The potential of selectively targeting advertisements on a user-by-user basis has been unrealized due to the difficulty in performing meaningful targeting of customers. The current generation of web ad selection engines utilize a number of partially successful techniques to target customers. Typically, the effectiveness of these techniques is measured in terms of audience response rates. Audience response, also called "clickthrough", is evaluated by counting the number of users that click on a "banner advertisement" contained on a web page which is presented to the user. Clicking on the banner advertisement typically takes the user to the web site of the advertiser where additional information about a product or service is provided. In the current Internet advertising environment, clickthrough is the best measure of the effectiveness of advertising techniques. Consequently, the value of advertising is directly related to the effectiveness of the ad. Therefore, the maximization of clickthrough is of paramount importance for most web sites for both practical and financial reasons.

Brief Summary Text (6):

Some current banner advertising selection techniques are listed in the tables below. These techniques are divided into two classes. The first class of advertising selection techniques, shown in Table 1, use simplistic first generation techniques that are based only on static a priori information. The second class of advertising selection techniques, shown in Table 2, utilizes some of the techniques used in the

more sophisticated second generation ad selection systems which may take into account some measure of user behavior, such as a user query, or make use of predisclosed user preferences. However, these more sophisticated objectives often only complicate the problem. In Table 1 and Table 2 certain disadvantages of each method are given, however, the Tables are not meant to be all inclusive so not all disadvantages may be shown.

Brief Summary Text (7) :

In the current web environment, users utilize search services to find relevant or interesting information. These search services provide a potential focus for the identification of user behavior as defined by the searches they execute, the web pages they view and positions in the directory hierarchy they visit. However, existing banner advertising selection techniques fail to analyze this behavior when selecting an advertisement. These search services provide an opportunity for presentation of user specific advertising.

Brief Summary Text (18) :

Rule based ad selection requires extensive knowledge about the targeted operating domain. Even with computer-aided tools, a knowledge engineer is required to develop the rules and administer the system. Furthermore, humans have demonstrated time and again, that they are poor at encoding rules. This observation is particularly true when large numbers of variables are encompassed within the scope of the problem being modeled.

Brief Summary Text (20) :

In keyword based ad selection systems, the ads are selected on the basis of one or more user provided words. When an observed user behavior (typically a user issued query) contains a known keyword, one of the ads, which is manually associated with the keyword, is selected for display. This technique provides good response rates for the keywords chosen. However, a major drawback with this approach is the system administrator must manually choose the keywords associated with each ad. This technique, based on intellectual effort and deep knowledge of the ad-specific domain, is time consuming and error prone. Additionally, the "inventory" of keywords at a site quickly becomes sold out. Lastly, with keyword based ad selection techniques, the ad selection process does not account for previous user behavior. The ad selector only uses a set of human-selected keywords in the current inventory based upon the current search query. Synonyms of user provided words are not automatically targeted without a thesaurus or synonym list.

Brief Summary Text (22) :

Ads are sometimes manually targeted to individual web pages. This method requires human intellectual effort to match an advertisement to a web page. However, such effort is usually prohibitive in large scale sites containing thousands of web pages.

Brief Summary Text (24) :

Ads are sometimes manually targeted to hierarchies or categorizations of pages called page ontologies. This technique may provide effective performance if the human making the assignment of the ad to a branch of the hierarchy has a good feel for the content of the site and understanding of the intended viewer. However, human intellectual effort is required to select the appropriate branch or category for the ad.

Brief Summary Text (26) :

Ads are sometimes manually targeted to all of the pages in a web site. However, as was seen with many of the other systems, human intellectual effort is required to select the appropriate ad for the web site. This method is inflexible in that an advertiser cannot automatically target an ad to the best pages within a site. Typically, advertisers want to display different ads for each page in a site. Another problem with assigning ads to web sites is that the ads served to the user are not context dependent in that they fail to utilize user specific information.

Brief Summary Text (30) :

The existing ad selection systems, particularly in coupon and print media advertising, fail to address the interactive nature of the Internet and electronic commerce. Advertisers need to be able to identify users of specific interests, track those interests over time and disseminate, in a highly selective way, information,

advertising, coupons and product offerings that will be of interest to the user. Additionally, advertisers need to track user interests and behavior in a real-time manner.

Brief Summary Text (31):

Therefore, advertisers need a system which is sensitive to user behavior for targeting advertising, coupons, products and information content. This system should enable the targeted marketing in real-time and to a granularity of an individual user as well as groups of users that have similar behavioral characteristics.

Brief Summary Text (33):

The present invention includes a system and method for delivering targeted entities, such as advertising, coupons, products and information content, to users based on tracking observed behavior. The system model is an adaptive vector space representation for entities, information and

Brief Summary Text (36):

Web pages, banner advertisements, coupons, and any other textual or symbolic information, are represented as a summary content vector by forming the normalized weighted vector sum of the content vectors of the words (symbols) in the page or description of the product or service. These summary content vectors have the property that pages that discuss similar information have summary content vectors that point in similar directions.

Brief Summary Text (39):

The present system also performs the optimal adaptive positioning of entity vectors. If a user is presented a banner ad and requests more information by clicking the mouse on it, the entity vector for the ad is adjusted using the user's profile vector. Additionally, the user profile vector is adjusted using the ad content vector. This mutual adaptation allows optimum positioning of ads relative to users.

Detailed Description Text (6):

In one implementation, the server 106 includes a gateway which is connected to a WAN 108. The WAN 108 has a plurality of network servers 110. One of the network servers 110 is connected to a LAN 112 comprising a plurality of computers 114. The system software that selects the advertisement may be located on one the network servers 110 or another computer in the network 104. In one embodiment of the invention, the advertisement selection software executes in part on a plurality of the network servers 110. In another embodiment of the invention, the advertisement selection software executes on a plurality of the computers 114 on the LAN 112. In yet another embodiment of the invention, the advertisement selection software resides on the client computer 100. It is important to understand that the advertisement selection system of the present invention may be hosted on any computing device so long as a communication pathway exists between the advertisement selection software and the requester of the advertisement FIG. 2 is a functional block diagram illustrating the processing and data components of the present invention. The following description focuses on the system's ability to deliver advertising, specifically banner advertising over the Internet. However, the techniques utilized in this invention are general in nature and can be applied to any type of entity or information. Furthermore, the network environment need not be the Internet as indicated previously.

Detailed Description Text (21):

The PVUS module 224 adapts the user profile vectors in response to a request from the RTAS module 214. A profile vector for each user is stored in a profile vector database 226. This adaptation process moves the user profile vector in the direction of the entity vector associated with the adaptation event. For instance, if the user clicks on an ad, the user profile vector is moved a small step in the direction of the entity vector for the ad.

Detailed Description Text (22):

The EVUS module 228 performs a nearly identical action as the PVUS module 224, but performs the operation on the entity vectors for advertisements. Each entity vector is stored in an entity vector database 230. For example, if the user clicks on an ad, the entity vector for the user selected ad is moved a small step in the direction of the

user's profile vector.

Detailed Description Text (24) :

It should be noted that the adaptation of the entity vectors is not required for system operation. Indeed, when adaptation is enabled, it allows the entity vector of the product to "drift" toward users who have received the ad. In some applications, information content delivery systems in particular, this drift is not desirable. Clearly, the information content of a news story is deterministic and should not change based on who reads the story. As such, the present invention has provisions to disable the adaptation of entity vectors. However, in many applications, such as Internet advertising, the drift of the vectors is the mechanism whereby the system discovers patterns and trends.

Detailed Description Text (82) :

Targeted impression count. The targeted impression count determines the number of impressions that should be delivered for the ad.

Detailed Description Text (117) :

(i) dissemination of targeted advertising, coupons, offers and information to the user;

Detailed Description Text (119) :

(iii) correlation of user behavior with advertising, product and service descriptions;

Detailed Description Text (123) :

(vii) accumulation of effectiveness and click-through statistics to optimize predictive models, provide management reports and lay groundwork for increased ad revenues.

Detailed Description Text (134) :

If the user clicks on an ad or purchases a product 222, the TACV module 1704 uses the ad or product description to adapt the user profile vector. This module 1704 transforms the entity ID into a content vector via a table lookup in memory within the PVUS module 224. In parallel, the user ID is used to lookup the user profile vector from the profile vector database 226. The PVA module 1702 inputs these two vectors into the adaptation algorithm described in FIGS. 19a and 19b. The profile vector database update module 1704 then writes updated profile vector to the profile vector database 226 for later use.

Detailed Description Text (136) :

For many advertising and sales applications, it is advantageous to have an estimate of the demographic parameters for individuals as well as the entire user population. In the present system, although no user registration is performed, advertisers can analyze the profiles vectors to retrieve user demographics. Advertisers can use these profile vectors to estimate the age, gender, household income (HHI), and other demographic parameters of a user. The present system has provisions to incorporate demographic models into its operation. When using these demographic models, the demographic parameters of the entire population of users are estimated and used to augment the information in the profile vector database 226. This estimate process can be performed by both conventional statistical methods or via neural network techniques.

Detailed Description Text (149) :

An adaptation event occurs when a user enters a query 217 or clicks on an ad or performs an on-line purchase of a product 222. When the entity adaptation event occurs, the user action that resulted in the event, such as a user query 217, the EVUS module 228 obtains from the RTAS module 214 the entity ID for the selected ad and the user ID.

Detailed Description Text (155) :

The EVUS module 228 performs inventory management weighting across all the ads in the system. The EVUS module 228 uses a weight parameter to adjust each of the ad selection scores. The weight parameter is determined by the ads targeted number of impressions and the time remaining for each ad.

Detailed Description Text (171) :

The RTAS module 214 keeps track of the number of impressions generated for each entity or ad and periodically sends a message to the EVUS module 228 containing each ad ID and the number of impressions generated for each ad. The RTAS module 214 uses the ad parameters specifying the targeted number of impressions and the time each ad will remain eligible for serving to calculate inventory management weights. The weights are applied by the RTAS module 214 when scoring the active ads. The RTAS module 214 causes under-impressed ads to receive more impressions while suppressing impressions for over-impressed ads.

Detailed Description Text (177) :

The RDA module 236 performs unsupervised categorization or clustering of the contents of the profile vector database 225, the entity vector database 230 or both of these modules. The RDA module 236 may use any clustering algorithm as long as the algorithm can process real number vectors. The result of the clustering is a set of cluster centroid vectors that represent regions of high density. These centroids are the same type of vectors that are used for all the other processing operations. Thus, the RDA system is able to find a set of word vectors that are closest in the content vector space to the centroid vector. The words associated with these vectors, and their closeness metrics, constitute the key concepts of the centroid. The evolution of cluster centroids over time provides trend information. The identification of word vectors close to these cluster centroids provides a key word summary for each cluster, easing the analysis and interpretation task. Conventional statistical analysis may also be used to investigate cluster contents. The variance of cluster membership is a reliable indication of cluster coherence and "unity of focus" in terms of behavior. The cluster identified in the product vector database 226 provides highly valuable information about the characteristics of the users who purchase the products. This result is a natural consequence of the mutual adaptation between the ad vectors and the user content vectors. The number of clusters found and the explanation of the clusters are listed in the analysis reports 238 for inspection by the analyst 240.

Detailed Description Text (182) :

A summary of the operation of the system for the advertising scenario is as follows.

Detailed Description Text (184) :

First, the word content vector learning module 200 creates multiple content vectors from text input. This technique, results in a set of word content vectors, one word content vector for each word found in the training vocabulary and not in the stop list 204. In the advertising scenario, the training texts are usually product descriptions and trade magazines. For instance, if the targeted domain of operation is sporting goods, the training corpus would contain the product descriptions from sporting goods catalogs, the contents of several years of "Field and Stream" magazine and books on outdoor life. The selection of the correct training corpus, oriented toward the domain of operation, such as outdoor sports, gives the correct "spin" on the usage of words. In the aforementioned example, "bow" becomes associated with "arrow" as opposed to "violin" or an action performed on stage at the completion of a performance.

Detailed Description Text (193) :

During discovery and analysis mode, at periodic intervals, the system administrator performs relationship discovery and analysis of the contents of both the entity vector database 230 and the user profile vector database 226 to discover useful and exploitable characteristics of user behavior. The RDA module 236 performs vector clustering, reporting and provides summary statistical information of system effectiveness in terms of user clickthrough rate. The behavior clusters discovered during analysis serves as the basis of advertising campaigns and provides valuable insight into user behavioral preferences. The behavior clusters can be utilized as the initial conditions for entity vectors the advertiser wants to target by group, rather than on an individual basis.